Membership problem in $\mathrm{GL}(2,\mathbb{Z})$ extended by singular matrices*

Igor Potapov[1] and Pavel Semukhin[2]

1 **Department of Computer Science, University of Liverpool, United Kingdom**
potapov@liverpool.ac.uk
2 **Department of Computer Science, University of Liverpool, United Kingdom**
semukhin@liverpool.ac.uk

Abstract

We consider the membership problem for matrix semigroups, which is the problem to decide whether a matrix belongs to a given finitely generated matrix semigroup.

In general, the decidability and complexity of this problem for two-dimensional matrix semigroups remains open. Recently there was a significant progress with this open problem by showing that the membership is decidable for 2×2 nonsingular integer matrices. In this paper we focus on the membership for singular integer matrices and prove that this problem is decidable for 2×2 integer matrices whose determinants are equal to 0, 1, -1 (i.e. for matrices from $\mathrm{GL}(2,\mathbb{Z})$ and any singular matrices). Our algorithm relies on a translation of numerical problems on matrices into combinatorial problems on words and conversion of the membership problem into decision problem on regular languages.

1998 ACM Subject Classification F.2.1 Numerical Algorithms and Problems, F.1.1 Models of Computation

Keywords and phrases Matrix Semigroups, Membership Problem, General Linear Group, Singular Matrices, Automata and Formal Languages

Digital Object Identifier 10.4230/LIPIcs.MFCS.2017.44

1 Introduction

Matrices and matrix products play a crucial role in the representation and analysis of various computational processes such as linear recurrent sequences [13, 21, 22], arithmetic circuits [11], hybrid and dynamical systems [20, 2], probabilistic and quantum automata [7], stochastic games, broadcast protocols [10]. Many problems for matrices in dimension three and four are undecidable, but the decidability and complexity of problems for two-dimensional matrix semigroups remains open. One of such hard questions is the Membership problem.

Membership problem: Given a finite set of $m \times m$ matrices $\mathcal{F} = \{M_1, M_2, \ldots, M_n\}$ and a matrix M. Determine whether there exist an integer $k \geq 1$ and $i_1, i_2, \ldots, i_k \in \{1, \ldots, n\}$ such that

$$M_{i_1} \cdot M_{i_2} \cdots M_{i_k} = M.$$

In other words, determine whether a matrix M belongs to the semigroup generated by $\mathcal{F}$.

There are only few known decidability results for the membership problem when the dimension is not bounded. In 1986 Kannan and Lipton [14] proved that the membership is

* This work was supported by EPSRC grant "Reachability problems for words, matrices and maps" (EP/M00077X/1).

decidable in polynomial time for a semigroup generated by a single $m \times m$ matrix (known as the *Orbit problem*). Later, in 1996 this decidability result was extended to a more general case of commutative matrices [1]. A generalization of this result to a special class of non-commutative matrices (a class of row-monomial matrices over a commutative semigroup satisfying some natural effectiveness conditions) was shown in 2004 in [16]. On the other hand, it is known that the membership is already undecidable for 3×3 integer matrices with determinant 0 (i.e. singular matrices), see [23]. As for the nonsingular case, it is known that the membership is undecidable for 4×4 integer matrices with determinant one, see [5]. A more recent survey of undecidable problems can be found in [8].

Due to a severe lack of methods and techniques the status of decision problems for 2×2 matrices (like membership, vector reachability, freeness) remains a long standing open problem. Recently, a new approach of translating numerical problems on 2×2 integer matrices into a variety of combinatorial and computational problems on words over group alphabet and studying their transformations as specific rewriting systems have led to new results on decidability and complexity. The main ingredient of the translation into combinatorial problems on words is the well-known result that the groups $\mathrm{SL}(2,\mathbb{Z})$ and $\mathrm{GL}(2,\mathbb{Z})$ are finitely generated. For example, $\mathrm{SL}(2,\mathbb{Z})$ can be generated by a pair of matrices

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

with the following relations: $S^4 = I$, $R^6 = I$ and $S^2 = R^3$. So, we can represent a matrix $M \in \mathrm{SL}(2,\mathbb{Z})$ as a word in the alphabet $\{S, R\}$.

In particular, this symbolic representation was successfully used to show the decidability of the membership problem for the semigroups of $\mathrm{GL}(2,\mathbb{Z})$ [9] in 2005 and of the mortality problem for 2×2 integer matrices with determinants $0, \pm 1$ [19] in 2008. It also found applications in the design of the polynomial time algorithm for the membership problem for the modular group [12] in 2007. Furthermore, it was used to show NP-hardness for most of the reachability problems in dimension two [6, 3] in 2012 and to prove decidability of the vector/scalar reachability problems in $\mathrm{SL}(2,\mathbb{Z})$ [24] in 2016.

In 2017 a significant progress was made towards decidability of the membership problem for 2×2 integer matrices extending previously known result on $\mathrm{GL}(2,\mathbb{Z})$ [9]. In [25] the first algorithm was discovered that can check the membership problem for a matrix semigroup generated by nonsingular 2×2 integer matrices. In this paper we show another extension of [9] and prove that the membership problem is decidable for 2×2 integer matrices whose determinants are equal to 0, 1, -1 (i.e. for matrices from $\mathrm{GL}(2,\mathbb{Z})$ and any singular matrices). As a first step we give an alternative proof of the decidability of the mortality problem (i.e. membership for the zero matrix) from [19], in which we will use the Smith normal forms of matrices. In contrast to [19], our new approach allows us to generalize this proof to show decidability of the membership problem for singular matrices. The algorithm is based on a nontrivial combination of algebraic properties of $\mathrm{GL}(2,\mathbb{Z})$, automata theory and properties of matrix products with singular matrices.

2 Preliminaries

The semigroup of 2×2 integer matrices is denoted by $\mathbb{Z}^{2\times 2}$. Let $\mathrm{GL}(2,\mathbb{Z})$ be the general linear group of dimension 2 over $\mathbb{Z}$, that is, the group of 2×2 integer matrices whose determinant is equal to ± 1. We will use $\mathbf{O}$ to denote the zero 2×2 matrix. A matrix is called *singular* if its determinant is equal to zero and *nonsingular* otherwise.

If $\mathcal{F}$ is a finite collection of matrices from $\mathbb{Z}^{2\times 2}$, then $\langle\mathcal{F}\rangle$ denotes the semigroup generated by $\mathcal{F}$ (including the identity matrix), that is, $M \in \langle\mathcal{F}\rangle$ if and only if $M = I$ or there are matrices $M_1, \ldots, M_n \in \mathcal{F}$ such that $M = M_1 \cdots M_n$.

Consider the following matrices from $\mathrm{GL}(2,\mathbb{Z})$:

$$N = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},\ S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ R = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix},\ X = -I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Every word in the alphabet $\Sigma = \{N, S, R, X\}$ corresponds to a matrix from $\mathrm{GL}(2,\mathbb{Z})$ in a natural way. Namely, the letters N, S, R, X correspond to the matrices defined by the above formulas, and a word $w \in \Sigma^*$ corresponds to the product of its letters. For example, the word SR corresponds to the matrix $\begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}$.

It is well-known that $\mathrm{GL}(2,\mathbb{Z})$ is generated by the above matrices. So any $M \in \mathrm{GL}(2,\mathbb{Z})$ can be presented by a word in the alphabet $\Sigma = \{N, S, R, X\}$. Such presentation is not unique because of the identities like $S^2 = R^3 = X$. However for every matrix $M \in \mathrm{GL}(2,\mathbb{Z})$, there is a unique canonical word that represents it, as described below.

▶ **Definition 1.** A word $w \in \Sigma^*$ is called a *canonical word* if it has the form

$$w = N^\delta X^\gamma S^\beta R^{\alpha_0} S R^{\alpha_1} S R^{\alpha_2} \ldots S R^{\alpha_{n-1}} S R^{\alpha_n},$$

where $\beta, \delta, \gamma \in \{0, 1\}$, $\alpha_0, \ldots, \alpha_{n-1} \in \{1, 2\}$, and $\alpha_n \in \{0, 1, 2\}$. In other words, w is *canonical* if it does not contain subwords SS or RRR. Moreover, letter N may appear only once in the first position, and letter X may appear only once either in the first position or after N.

▶ **Proposition 2** ([17, 18, 26, 25])**.** *For every $M \in \mathrm{GL}(2,\mathbb{Z})$, there is a unique canonical word w which represents M.*

We will also use two additional matrices $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and the following identities: $R = ST$, $T = XSR$, $T^{-1} = XR^2S$, $U = XSR^2$ and $U^{-1} = XRS$.

▶ **Definition 3.** A subset $\mathcal{S} \subseteq \mathrm{GL}(2,\mathbb{Z})$ is called *regular* or *automatic* if there is a regular language L in alphabet Σ that describes $\mathcal{S}$. That is, every word $w \in L$ corresponds to a matrix M from $\mathcal{S}$, and for every matrix $M \in \mathcal{S}$, there is a word $w \in L$ that represents M.

▶ **Definition 4.** We call two words w_1 and w_2 from Σ^* *equivalent*, denoted $w_1 \sim w_2$, if they represent the same matrix. Two languages L_1 and L_2 in the alphabet Σ are *equivalent*, denoted $L_1 \sim L_2$, if
(i) for each $w_1 \in L_1$, there exists $w_2 \in L_2$ such that $w_1 \sim w_2$, and
(ii) for each $w_2 \in L_2$, there exists $w_1 \in L_1$ such that $w_2 \sim w_1$.
In other words, L_1 and L_2 are equivalent if and only if they describe the same language. Two finite automata $\mathcal{A}_1$ and $\mathcal{A}_2$ over alphabet Σ are *equivalent*, denoted $\mathcal{A}_1 \sim \mathcal{A}_2$, if $L(\mathcal{A}_1) \sim L(\mathcal{A}_2)$.

The next theorem will be a crucial ingredient of our decidability result.

▶ **Theorem 5.** *Given two regular subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathrm{GL}(2,\mathbb{Z})$, it is decidable whether the intersection $\mathcal{S}_1 \cap \mathcal{S}_2$ is empty or not.*

Proof. The proof is based on the following result: for every finite automaton $\mathcal{A}$ over alphabet Σ, there is an automaton $\mathrm{Can}(\mathcal{A})$ such that $\mathrm{Can}(\mathcal{A})$ accepts only canonical words and $\mathrm{Can}(\mathcal{A}) \sim \mathcal{A}$, that is, $\mathrm{Can}(\mathcal{A})$ and $\mathcal{A}$ describe the same subset of $\mathrm{GL}(2,\mathbb{Z})$. The construction of $\mathrm{Can}(\mathcal{A})$ can be found in [25], see also [9] for an alternative construction.

Now let L_1 and L_2 be regular languages that describe $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, and let $\mathcal{A}_1$ and $\mathcal{A}_2$ be finite automata such that $L(\mathcal{A}_1) = L_1$ and $L(\mathcal{A}_2) = L_2$. By Proposition 2, every matrix from $\mathrm{GL}(2,\mathbb{Z})$ corresponds to a unique canonical word. Therefore, we obtain the following equivalence: $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$ if and only if the languages of the automata $\mathrm{Can}(\mathcal{A}_1)$ and $\mathrm{Can}(\mathcal{A}_2)$ have nonempty intersection. Since the emptiness problem for regular languages is decidable, it is decidable whether $\mathcal{S}_1 \cap \mathcal{S}_2$ is empty or not. ◀

Another important ingredient of our proof is the existence and uniqueness of the Smith normal form of a matrix.

▶ **Theorem 6** (Smith normal form [15]). *For any matrix $A \in \mathbb{Z}^{2\times 2}$, there are matrices E, F from $\mathrm{GL}(2,\mathbb{Z})$ such that $A = E \begin{bmatrix} t_1 & 0 \\ 0 & t_2 \end{bmatrix} F$ for some nonnegative integers t_1 and t_2 such that $t_1 \mid t_2$. The diagonal matrix $\begin{bmatrix} t_1 & 0 \\ 0 & t_2 \end{bmatrix}$, which is unique, is called the* Smith normal form *of A. In fact, t_1 is equal to the gcd of the coefficients of A. Moreover, E, F, t_1, and t_2 can be computed in polynomial time.*

Remark. If $A \in \mathbb{Z}^{2\times 2}$ is nonzero matrix with $\det(A) = 0$, then the Smith normal form of A is equal to $\begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}$, where t is the gcd of the coefficients of A.

Using uniqueness of the Smith normal form we obtain the following corollary.

▶ **Corollary 7.** *If $E, F \in \mathrm{GL}(2,\mathbb{Z})$, then $E \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F$ is not a zero matrix.*

3 Main result

The main result of our paper is the following theorem.

▶ **Theorem 8.** *Let $M \in \mathbb{Z}^{2\times 2}$ and let $\mathcal{F} = \{A_1, \ldots, A_n, B_1, \ldots, B_m\}$ be a collection of matrices from $\mathbb{Z}^{2\times 2}$ such that $A_i \in \mathrm{GL}(2,\mathbb{Z})$ for $i = 1, \ldots, n$, and B_j is a singular matrix for $j = 1, \ldots, m$. Then it is decidable whether $M \in \langle \mathcal{F} \rangle$.*

Proof. First, note that if $M \in \langle \mathcal{F} \rangle$, then M is either singular or $M \in \mathrm{GL}(2,\mathbb{Z})$. Therefore, if $|\det(M)| > 1$, then we know that $M \notin \langle \mathcal{F} \rangle$. On the other hand, if $\det(M) = \pm 1$, i.e. if $M \in \mathrm{GL}(2,\mathbb{Z})$, then $M \in \langle \mathcal{F} \rangle$ if and only if $M \in \langle A_1, \ldots, A_n \rangle$. In other words, our problem reduces to the membership problem in $\mathrm{GL}(2,\mathbb{Z})$, and the decidability of the membership in $\mathrm{GL}(2,\mathbb{Z})$ was proven in [9].

Hence from now on we will assume that M is a singular matrix. First, we consider the case when M is the zero matrix and after that we consider the case when M is a nonzero singular matrix. The case when $M = \mathbf{O}$ is also called the *Mortality problem*. The decidability of the mortality problem for matrices with determinants 0 and ± 1 was shown in [19]. In Theorem 11 below we provide an alternative proof of this fact, which is based on the use of the Smith normal form of a matrix. Another reason why we include the proof of Theorem 11 is that it presents a simplified version of a more complicated construction for a nonzero M. Finally, in Theorem 13 we prove decidability of the membership problem for a nonzero singular matrix M. ◀

First, we prove Proposition 9 which will play an important role in the proofs of Theorems 11 and 13. Moreover, Proposition 9 and Corollary 10 reveal new structural properties of certain subsets of $\mathrm{GL}(2,\mathbb{Z})$ in symbolic presentation.

▶ **Proposition 9.** *For any fixed $a \in \mathbb{Z}$, let $M(a) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ b,c,d \in \mathbb{Z} \right\}$. Then $M(a)$ is a regular subset of $\mathrm{GL}(2,\mathbb{Z})$.*

Proof. First, suppose that $a = 0$. Then

$$M(0) = \left\{ \begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix} : d \in \mathbb{Z} \right\} \bigcup \left\{ \begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix} : d \in \mathbb{Z} \right\} \bigcup \left\{ \begin{bmatrix} 0 & 1 \\ 1 & d \end{bmatrix} : d \in \mathbb{Z} \right\} \bigcup \left\{ \begin{bmatrix} 0 & -1 \\ -1 & d \end{bmatrix} : d \in \mathbb{Z} \right\}.$$

Note that $\begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix} = ST^d$, $\begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix} = -ST^{-d}$, $\begin{bmatrix} 0 & 1 \\ 1 & d \end{bmatrix} = SNT^d$, $\begin{bmatrix} 0 & -1 \\ -1 & d \end{bmatrix} = -SNT^{-d}$. Hence we can express $M(0)$ as

$$\begin{aligned} M(0) =& \{ST^d : d \in \mathbb{Z}\} \cup \{-ST^{-d} : d \in \mathbb{Z}\} \cup \{SNT^d : d \in \mathbb{Z}\} \cup \{-SNT^{-d} : d \in \mathbb{Z}\} = \\ & \{ST^d : d \geq 0\} \cup \{S(T^{-1})^d : d \geq 0\} \cup \{-S(T^{-1})^d : d \geq 0\} \cup \\ & \{-ST^d : d \geq 0\} \cup \{SNT^d : d \geq 0\} \cup \{SN(T^{-1})^d : d \geq 0\} \cup \\ & \{-SN(T^{-1})^d : d \geq 0\} \cup \{-SNT^d : d \geq 0\}. \end{aligned}$$

Therefore, $M(0)$ can be described by the following regular expression

$$\begin{aligned} & S(XSR)^* + S(XR^2S)^* + XS(XR^2S)^* + XS(XSR)^* + \\ & SN(XSR)^* + SN(XR^2S)^* + XSN(XR^2S)^* + XSN(XSR)^*. \end{aligned}$$

Now suppose that $a \neq 0$. Consider a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z})$ and let $b = b_0 + ma$ and $c = c_0 + na$, where $b_0, c_0 \in \{0, \ldots, |a| - 1\}$ and $m, n \in \mathbb{Z}$. Since $A \in \mathrm{GL}(2,\mathbb{Z})$, we have $ad - bc = \pm 1$ or

$$d = \frac{bc \pm 1}{a} = \frac{(b_0 + ma)(c_0 + na) \pm 1}{a} = \frac{b_0c_0 \pm 1}{a} + mc_0 + nb_0 + mna.$$

Since d is an integer, $\frac{b_0c_0 \pm 1}{a}$ must also be an integer. Note that

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b_0 + ma \\ c_0 + na & \frac{b_0c_0 \pm 1}{a} + mc_0 + nb_0 + mna \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ n & 1 \end{bmatrix} \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0 \pm 1}{a} \end{bmatrix} \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}.$$

So, $A = U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0 \pm 1}{a} \end{bmatrix} T^m$. Let $N^+(a)$ and $N^-(a)$ be the following finite sets

$$\begin{aligned} N^+(a) &= \{(b_0, c_0) \ : \ b_0, c_0 \in \{0, \ldots, |a| - 1\} \text{ and } \tfrac{b_0c_0 + 1}{a} \text{ is an integer}\}, \\ N^-(a) &= \{(b_0, c_0) \ : \ b_0, c_0 \in \{0, \ldots, |a| - 1\} \text{ and } \tfrac{b_0c_0 - 1}{a} \text{ is an integer}\}. \end{aligned}$$

Then

$$M(a) = \bigcup_{(b_0,c_0) \in N^+(a)} \left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} T^m \ : \ n,m \in \mathbb{Z} \right\} \bigcup \bigcup_{(b_0,c_0) \in N^-(a)} \left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0-1}{a} \end{bmatrix} T^m \ : \ n,m \in \mathbb{Z} \right\}.$$

For each $(b_0, c_0) \in N^+(a)$, let $w^+(b_0, c_0)$ be a word that represents the matrix $\begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix}$. Note that for every $(b_0, c_0) \in N^+(a)$, we can present $\left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} T^m : n, m \in \mathbb{Z} \right\}$ as a union of four sets

$$\left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} T^m : \ n, m \geq 0 \right\} \bigcup \left\{ (U^{-1})^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} T^m : \ n, m \geq 0 \right\} \bigcup$$
$$\left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} (T^{-1})^m : \ n, m \geq 0 \right\} \bigcup \left\{ (U^{-1})^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0+1}{a} \end{bmatrix} (T^{-1})^m : \ n, m \geq 0 \right\}.$$

Hence it can be described by the following regular expression

$$(XSR^2)^* w^+(b_0, c_0)(XSR)^* + (XRS)^* w^+(b_0, c_0)(XSR)^* +$$
$$(XSR^2)^* w^+(b_0, c_0)(XR^2S)^* + (XRS)^* w^+(b_0, c_0)(XR^2S)^*.$$

Similarly, we have that for every $(b_0, c_0) \in N^-(a)$, the set $\left\{ U^n \begin{bmatrix} a & b_0 \\ c_0 & \frac{b_0c_0-1}{a} \end{bmatrix} T^m : n, m \in \mathbb{Z} \right\}$ can be described by a regular expression. Since $M(a)$ is equal to a finite union of such sets, we conclude that it is also regular. ◀

▶ **Corollary 10.** *For every $i = 1, 2, 3, 4$ and any fixed $a \in \mathbb{Z}$, the following subset of $\mathrm{GL}(2,\mathbb{Z})$ is a regular set:* $M_i(a) = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ a_i = a \text{ and } a_j \in \mathbb{Z} \text{ for } j \neq i \right\}$.

Proof. By definition, $M_1(a) = M(a)$, hence it is regular by Proposition 9. Now let $L(a)$ be a regular language that describes $M(a)$ and let $K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. It is not hard to see that $M_2(a) = M(a) \cdot K$, $M_3(a) = K \cdot M(a)$ and $M_4(a) = K \cdot M(a) \cdot K$. Note that matrix K corresponds to the word NXS. Therefore, $M_2(a)$, $M_3(a)$ and $M_4(a)$ can be described by the regular languages $L(a) \cdot \{NXS\}$, $\{NXS\} \cdot L(a)$ and $\{NXS\} \cdot L(a) \cdot \{NXS\}$, respectively. ◀

3.1 Mortality problem

In this section we will give an alternative proof of the decidability of the mortality problem from [19] which will be based on the use of the Smith normal form of a matrix.

▶ **Theorem 11.** *The mortality problem for 2×2 integer matrices with determinants $0, \pm 1$ is decidable.*

This theorem will be a consequence of Theorem 5 and Propositions 9 and 12.

▶ **Proposition 12.** *Let $\mathcal{F} = \{A_1, \ldots, A_n, B_1, \ldots, B_m\}$ be a collection of matrices from $\mathbb{Z}^{2\times 2}$ such that $A_i \in \mathrm{GL}(2,\mathbb{Z})$ for $i = 1, \ldots, n$, and $\det(B_j) = 0$ for $j = 1, \ldots, m$. Then $\mathbf{O} \in \langle \mathcal{F} \rangle$ if and only if $B_j = \mathbf{O}$ for some j or there are indices $i, j \in \{1, \ldots, m\}$ and a matrix $C \in \langle A_1, \ldots, A_n \rangle$ such that $B_i C B_j = \mathbf{O}$.*

Proof. Suppose that $B_j \neq \mathbf{O}$ for every j. Under this assumption, if $\mathbf{O} \in \langle \mathcal{F} \rangle$ then there are indices $i_1, \ldots, i_s \in \{1, \ldots, m\}$ and matrices $C_1, \ldots, C_{s+1} \in \langle A_1, \ldots, A_n \rangle$ such that

$$C_1 B_{i_1} C_2 B_{i_2} \cdots C_s B_{i_s} C_{s+1} = \mathbf{O}. \tag{1}$$

By Theorem 6, we can write each matrix B_{i_r}, for $i = 1, \ldots, s$, as $B_{i_r} = E_r \begin{bmatrix} t_r & 0 \\ 0 & 0 \end{bmatrix} F_r$, where $E_r, F_r \in \mathrm{GL}(2, \mathbb{Z})$ and $t_r > 0$. Then (1) is equivalent to

$$C_1 E_1 \begin{bmatrix} t_1 & 0 \\ 0 & 0 \end{bmatrix} F_1 C_2 E_2 \begin{bmatrix} t_2 & 0 \\ 0 & 0 \end{bmatrix} F_2 \cdots \\ \cdots C_{s-1} E_{s-1} \begin{bmatrix} t_{s-1} & 0 \\ 0 & 0 \end{bmatrix} F_{s-1} C_s E_s \begin{bmatrix} t_s & 0 \\ 0 & 0 \end{bmatrix} F_s C_{s+1} = \mathbf{O}. \quad (2)$$

Dividing (2) by the product $t_1 t_2 \cdots t_{s-1} t_s$, which is nonzero by our assumption, we obtain

$$C_1 E_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_1 C_2 E_2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_2 \cdots C_{s-1} E_{s-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_{s-1} C_s E_s \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_s C_{s+1} = \mathbf{O}. \quad (3)$$

Suppose for each $r = 2, \ldots, s$ the matrix $F_{r-1} C_r E_r$ have the form $F_{r-1} C_r E_r = \begin{bmatrix} a_r & b_r \\ c_r & d_r \end{bmatrix}$. Note that $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$. Therefore, (3) is equivalent to

$$C_1 E_1 \begin{bmatrix} a_2 a_3 \cdots a_s & 0 \\ 0 & 0 \end{bmatrix} F_s C_{s+1} = \mathbf{O}. \quad (4)$$

Suppose that $a_2 a_3 \cdots a_s \neq 0$. In this case (4) is equivalent to

$$C_1 E_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_s C_{s+1} = \mathbf{O},$$

where $C_1 E_1$ and $F_s C_{s+1}$ are matrices from $\mathrm{GL}(2, \mathbb{Z})$. This contradicts Corollary 7. Hence $a_2 a_3 \cdots a_s = 0$, and therefore there is $r \in \{2, \ldots, s\}$ such that $a_r = 0$. This implies that $B_{i_{r-1}} C_r B_{i_r} = \mathbf{O}$. Indeed,

$$B_{i_{r-1}} C_r B_{i_r} = E_{r-1} \begin{bmatrix} t_{r-1} & 0 \\ 0 & 0 \end{bmatrix} F_{r-1} C_r E_r \begin{bmatrix} t_r & 0 \\ 0 & 0 \end{bmatrix} F_r.$$

By assumption, $F_{r-1} C_r E_r = \begin{bmatrix} a_r & b_r \\ c_r & d_r \end{bmatrix}$. Thus

$$B_{i_{r-1}} C_r B_{i_r} = t_{r-1} t_r E_{r-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_r & b_r \\ c_r & d_r \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_r = t_{r-1} t_r E_{r-1} \begin{bmatrix} a_r & 0 \\ 0 & 0 \end{bmatrix} F_r.$$

Since $a_r = 0$, we have $B_{i_{r-1}} C_r B_{i_r} = \mathbf{O}$.

The implication on the other direction is trivial. ◀

Proof of Theorem 11. Obviously, if $B_j = \mathbf{O}$ for some $j = 1, \ldots, m$, then $\mathbf{O} \in \langle \mathcal{F} \rangle$. Therefore, from now on we assume that all B_j's are nonzero singular matrices. In this case Proposition 12 implies that $\mathbf{O} \in \langle \mathcal{F} \rangle$ if and only if there are indices $i, j \in \{1, \ldots, m\}$ and a matrix $C \in \langle A_1, \ldots, A_n \rangle$ such that $B_i C B_j = \mathbf{O}$. We now show that the latter property is algorithmically decidable.

Let $B_i = E_i \begin{bmatrix} t_i & 0 \\ 0 & 0 \end{bmatrix} F_i$ and $B_j = E_j \begin{bmatrix} t_j & 0 \\ 0 & 0 \end{bmatrix} F_j$ be the Smith normal forms of B_i and B_j, respectively. Note that by our assumption $t_i, t_j > 0$. Let C be a matrix from $\langle A_1, \ldots, A_n \rangle$. We have $B_i C B_j = \mathbf{O}$ if and only if

$$E_i \begin{bmatrix} t_i & 0 \\ 0 & 0 \end{bmatrix} F_i C E_j \begin{bmatrix} t_j & 0 \\ 0 & 0 \end{bmatrix} F_j = \mathbf{O} \quad \text{or, equivalently,} \quad E_i \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_i C E_j \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_j = \mathbf{O}.$$

Let $F_iCE_j = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for some $a, b, c, d \in \mathbb{Z}$. Then the above equation is equivalent to $E_i \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} F_j = \mathbf{O}$ or $E_i \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} F_j = \mathbf{O}$. By Corollary 7, $E_i \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} F_j = \mathbf{O}$ if and only if $a = 0$. So, we showed the following equivalence: $B_iCB_j = \mathbf{O}$ if and only if $F_iCE_j = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix}$ for some $b, c, d \in \mathbb{Z}$.

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be the following subsets of $\mathrm{GL}(2,\mathbb{Z})$: $\mathcal{S}_1 = \{F_iCE_j : C \in \langle A_1, \ldots, A_n \rangle\}$ and $\mathcal{S}_2 = \left\{ \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ b, c, d \in \mathbb{Z} \right\}$. In this notations the above equivalence can be written as follows: there is a matrix $C \in \langle A_1, \ldots, A_n \rangle$ such that $B_iCB_j = \mathbf{O}$ if and only if $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$.

It is easy to see that $\mathcal{S}_1$ is a regular subset of $\mathrm{GL}(2,\mathbb{Z})$ as it can be described by the regular expression $u_i(w_1 + \cdots + w_n)^* v_j$, where $w_1, \ldots, w_n$ are words representing the matrices $A_1, \ldots, A_n$ and u_i, v_j represent the matrices F_i, E_j, respectively. By Proposition 9, $\mathcal{S}_2$ is also a regular subset of $\mathrm{GL}(2,\mathbb{Z})$. Using Theorem 5, we can decide whether $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$ and hence decide whether there is a matrix $C \in \langle A_1, \ldots, A_n \rangle$ such that $B_iCB_j = \mathbf{O}$. This finishes the proof of Theorem 11. ◀

3.2 Membership problem

We are now ready to consider the case when M is a nonzero singular matrix.

▶ **Theorem 13.** *Let $\mathcal{F} = \{A_1, \ldots, A_n, B_1, \ldots, B_m\}$ be a finite collection of matrices such that $A_1, \ldots, A_n \in \mathrm{GL}(2,\mathbb{Z})$ and $B_1, \ldots, B_m$ are singular matrices from $\mathbb{Z}^{2\times 2}$. Also let $M \in \mathbb{Z}^{2\times 2}$ be a nonzero singular matrix. Then it is decidable whether $M \in \langle \mathcal{F} \rangle$.*

Proof. Let $M = E \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} F$ be the Smith normal form of M, and for each $j = 1, \ldots, m$, let $B_j = E_j \begin{bmatrix} t_j & 0 \\ 0 & 0 \end{bmatrix} F_j$ be the Smith normal form of B_j. Since M is a nonzero matrix, we may assume that for each $j = 1, \ldots, m$, B_j is also a nonzero matrix. Hence without loss of generality we assume that $t, t_1, \ldots, t_m > 0$.

We will construct a graph $\mathcal{G}(M, \mathcal{F})$, depending on M and $\mathcal{F}$, which will have the following property: $M \in \langle \mathcal{F} \rangle$ if and only if there is a path in $\mathcal{G}(M, \mathcal{F})$ from an initial to a final node of weight t.

Description of $\mathcal{G}(M, \mathcal{F})$. Graph $\mathcal{G}(M, \mathcal{F})$ has m nodes labelled by singular matrices $B_1, \ldots, B_m$ and two special nodes **In** and **Fin**, where **In** is the only initial node and **Fin** is the only final node. The weights of the nodes are defined as follows.

▶ **Definition 14.** Recall that $E_j \begin{bmatrix} t_j & 0 \\ 0 & 0 \end{bmatrix} F_j$ is the Smith normal form of B_j. Then the weight of the node with label B_j is equal to t_j.

Furthermore, we add edges to this graph according to the following rules.

▶ **Definition 15.** (1) For every integer $u \neq 0$ such that $-t \leq u \leq t$ we add an edge from node B_i to node B_j of weight u if and only if there is a matrix $C \in \langle A_1, \ldots, A_n \rangle$ such that $F_iCE_j \in \left\{ \begin{bmatrix} u & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ b, c, d \in \mathbb{Z} \right\}$.

(2) We also add an edge of weight u from the initial node **In** to a node with label B_j if there is a matrix $C \in \langle A_1, \dots, A_n \rangle$ such that $E^{-1}CE_j \in \left\{ \begin{bmatrix} u & b \\ 0 & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ b, d \in \mathbb{Z} \right\}$.

(3) Finally, we add an edge of weight u from a node with label B_j to the final node **Fin** if there is a matrix $C \in \langle A_1, \dots, A_n \rangle$ such that $F_jCF^{-1} \in \left\{ \begin{bmatrix} u & 0 \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ c, d \in \mathbb{Z} \right\}$.

Note that the set $\{F_iCE_j : C \in \langle A_1, \dots, A_n \rangle\}$ is a regular subset of $\mathrm{GL}(2,\mathbb{Z})$ because it can be described by the regular expression $u_i(w_1 + \cdots + w_n)^* v_j$, where $w_1, \dots, w_n$ are words representing the matrices $A_1, \dots, A_n$ and u_i, v_j represent the matrices F_i, E_j, respectively. By Proposition 9, $\left\{ \begin{bmatrix} u & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z}) \ : \ b, c, d \in \mathbb{Z} \right\}$ is also a regular subset of $\mathrm{GL}(2,\mathbb{Z})$. Therefore, by Theorem 5, we can algorithmically decide if there is an edge from node B_i to node B_j of weight u.

Moreover, the edges going out of **In** or ending in **Fin** can only have weights 1 or -1. Again, using Proposition 9, Corollary 10 and Theorem 5, we can algorithmically decide if there is an edge from **In** to B_j or from B_j to **Fin** of weight 1 or -1.

▶ **Definition 16.** The weight of a path in $\mathcal{G}(M, \mathcal{F})$ from **In** to **Fin** is equal to the product of the weights of nodes and edges that occur in it. That is, the weight of a path

$$\mathbf{In} \xrightarrow{u_0} B_{i_0} \xrightarrow{u_1} B_{i_1} \xrightarrow{u_2} B_{i_2} \quad \cdots \quad B_{i_{s-1}} \xrightarrow{u_s} B_{i_s} \xrightarrow{u_{s+1}} \mathbf{Fin}$$

is equal to $u_0 t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} u_{s+1}$.

In the following proposition we will show that the membership problem is equivalent to the existence of a path in $\mathcal{G}(M, F)$ with a given weight.

▶ **Proposition 17.** *Let* $M = E \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} F$ *be the Smith normal form of matrix* M. *Then* $M \in \langle \mathcal{F} \rangle$ *if and only if there is a path in* $\mathcal{G}(M, \mathcal{F})$ *from* **In** *to* **Fin** *of weight* t.

Proof. Suppose

$$\mathbf{In} \xrightarrow{u_0} B_{i_0} \xrightarrow{u_1} B_{i_1} \xrightarrow{u_2} B_{i_2} \quad \cdots \quad B_{i_{s-1}} \xrightarrow{u_s} B_{i_s} \xrightarrow{u_{s+1}} \mathbf{Fin}$$

is a path in $\mathcal{G}(M, \mathcal{F})$ from **In** to **Fin** of weight t. Recall that for every $r = 0, \dots, s$, we have $B_{i_r} = E_{i_r} \begin{bmatrix} t_{i_r} & 0 \\ 0 & 0 \end{bmatrix} F_{i_r}$. Hence $t = u_0 t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} u_{s+1}$.

Since for every $r = 1, \dots, s$ we have an edge $B_{i_{r-1}} \xrightarrow{u_r} B_{i_r}$ of weight u_r, there is a matrix $C_r \in \langle A_1, \dots, A_n \rangle$ such that $F_{i_{r-1}} C_r E_{i_r} = \begin{bmatrix} u_r & b_r \\ c_r & d_r \end{bmatrix}$ for some $b_r, c_r, d_r \in \mathbb{Z}$. Since we have an edge $\mathbf{In} \xrightarrow{u_0} B_{i_0}$ of weight u_0, there is a matrix $C_0 \in \langle A_1, \dots, A_n \rangle$ such that $E^{-1} C_0 E_{i_0} = \begin{bmatrix} u_0 & b_0 \\ 0 & d_0 \end{bmatrix}$ for some $b_0, d_0 \in \mathbb{Z}$. And since we have an edge $B_{i_s} \xrightarrow{u_{s+1}} \mathbf{Fin}$ of weight u_{s+1}, there is a matrix $C_{s+1} \in \langle A_1, \dots, A_n \rangle$ such that $F_{i_s} C_{s+1} F^{-1} = \begin{bmatrix} u_{s+1} & 0 \\ c_{s+1} & d_{s+1} \end{bmatrix}$ for some $c_{s+1}, d_{s+1} \in \mathbb{Z}$.

Hence we obtain the following equation

$$E^{-1} C_0 B_{i_0} C_1 B_{i_1} C_2 B_{i_2} \cdots B_{i_{s-1}} C_s B_{i_s} C_{s+1} F^{-1} =$$
$$E^{-1} C_0 E_{i_0} \begin{bmatrix} t_{i_0} & 0 \\ 0 & 0 \end{bmatrix} F_{i_0} C_1 E_{i_1} \begin{bmatrix} t_{i_1} & 0 \\ 0 & 0 \end{bmatrix} F_{i_1} C_2 E_{i_2} \begin{bmatrix} t_{i_2} & 0 \\ 0 & 0 \end{bmatrix} F_{i_2} \cdots$$

$$\cdots E_{i_{s-1}}\begin{bmatrix} t_{i_{s-1}} & 0 \\ 0 & 0 \end{bmatrix} F_{i_{s-1}} C_s E_{i_s} \begin{bmatrix} t_{i_s} & 0 \\ 0 & 0 \end{bmatrix} F_{i_s} C_{s+1} F^{-1} =$$

$$t_{i_0} t_{i_1} t_{i_2} \cdots t_{i_{s-1}} t_{i_s} \begin{bmatrix} u_0 & b_0 \\ 0 & d_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdots$$

$$\cdots \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_s & b_s \\ c_s & d_s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{s+1} & 0 \\ c_{s+1} & d_{s+1} \end{bmatrix} =$$

$$t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} \begin{bmatrix} u_0 & b_0 \\ 0 & d_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{s+1} & 0 \\ c_{s+1} & d_{s+1} \end{bmatrix} =$$

$$t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} \begin{bmatrix} u_0 u_{s+1} & 0 \\ 0 & 0 \end{bmatrix} = u_0 t_{i_0} u_1 t_{i_1} \cdots t_{i_{s-1}} u_s t_{i_s} u_{s+1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore, $C_0 B_{i_0} C_1 B_{i_1} C_2 B_{i_2} \cdots B_{i_{s-1}} C_s B_{i_s} C_{s+1} = E \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} F = M$, and hence $M \in \langle \mathcal{F} \rangle$.

Now suppose that $M \in \langle \mathcal{F} \rangle$. It is not hard to see that there is a sequence of indices $i_0, i_1, \ldots, i_s \in \{1, \ldots, m\}$, and matrices $C_0, C_1, \ldots, C_{s+1} \in \langle A_1, \ldots, A_n \rangle$ such that

$$C_0 B_{i_0} C_1 B_{i_1} C_2 B_{i_2} \cdots B_{i_{s-1}} C_s B_{i_s} C_{s+1} = M. \tag{5}$$

Recall that $E \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} F$ is the Smith normal form of M, and $E_{i_r} \begin{bmatrix} t_{i_r} & 0 \\ 0 & 0 \end{bmatrix} F_{i_r}$ is the Smith normal form of B_{i_r}, for $r = 0, \ldots, s$. So we can rewrite (5) as follows

$$\begin{aligned} E^{-1} C_0 E_{i_0} \begin{bmatrix} t_{i_0} & 0 \\ 0 & 0 \end{bmatrix} F_{i_0} C_1 E_{i_1} \begin{bmatrix} t_{i_1} & 0 \\ 0 & 0 \end{bmatrix} F_{i_1} C_2 E_{i_2} \begin{bmatrix} t_{i_2} & 0 \\ 0 & 0 \end{bmatrix} F_{i_2} \cdots \\ \cdots E_{i_{s-1}} \begin{bmatrix} t_{i_{s-1}} & 0 \\ 0 & 0 \end{bmatrix} F_{i_{s-1}} C_s E_{i_s} \begin{bmatrix} t_{i_s} & 0 \\ 0 & 0 \end{bmatrix} F_{i_s} C_{s+1} F^{-1} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{6}$$

For $r = 1, \ldots, s$, let $F_{i_{r-1}} C_r E_{i_r} = \begin{bmatrix} u_r & b_r \\ c_r & d_r \end{bmatrix}$. Then for every $r = 1, \ldots, s$, there is an edge $B_{i_{r-1}} \xrightarrow{u_r} B_{i_r}$ in $\mathcal{G}(M, \mathcal{F})$ of weight u_r. Furthermore, suppose that $E^{-1} C_0 E_{i_0} = \begin{bmatrix} u_0 & b_0 \\ c_0 & d_0 \end{bmatrix}$ and $F_{i_s} C_{s+1} F^{-1} = \begin{bmatrix} u_{s+1} & b_{s+1} \\ c_{s+1} & d_{s+1} \end{bmatrix}$. Then we can rewrite (6) as

$$\begin{bmatrix} u_0 & b_0 \\ c_0 & d_0 \end{bmatrix} \begin{bmatrix} t_{i_0} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} t_{i_1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} t_{i_2} & 0 \\ 0 & 0 \end{bmatrix} \cdots$$

$$\cdots \begin{bmatrix} t_{i_{s-1}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_s & b_s \\ c_s & d_s \end{bmatrix} \begin{bmatrix} t_{i_s} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{s+1} & b_{s+1} \\ c_{s+1} & d_{s+1} \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}$$

or equivalently

$$t_{i_0} t_{i_1} t_{i_2} \cdots t_{i_{s-1}} t_{i_s} \begin{bmatrix} u_0 & b_0 \\ c_0 & d_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdots$$

$$\cdots \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_s & b_s \\ c_s & d_s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{s+1} & b_{s+1} \\ c_{s+1} & d_{s+1} \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}.$$

From this equation we obtain

$$t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} \begin{bmatrix} u_0 & b_0 \\ c_0 & d_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{s+1} & b_{s+1} \\ c_{s+1} & d_{s+1} \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} \quad \text{or}$$

$$t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} \begin{bmatrix} u_0 u_{s+1} & u_0 b_{s+1} \\ c_0 u_{s+1} & c_0 b_{s+1} \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore, we have that $t = u_0 t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} u_{s+1}$ and $u_0 b_{s+1} = c_0 u_{s+1} = 0$. By assumption $t \neq 0$, and so $u_0 \neq 0$ and $u_{s+1} \neq 0$. Therefore, $c_0 = 0$ and $b_{s+1} = 0$. Hence we have that $E^{-1} C_0 E_{i_0} = \begin{bmatrix} u_0 & b_0 \\ 0 & d_0 \end{bmatrix}$ and $F_{i_s} C_{s+1} F^{-1} = \begin{bmatrix} u_{s+1} & 0 \\ c_{s+1} & d_{s+1} \end{bmatrix}$, which means that there is an edge $\mathbf{In} \xrightarrow{u_0} B_{i_0}$ of weight u_0 and an edge $B_{i_s} \xrightarrow{u_{s+1}} \mathbf{Fin}$ of weight u_{s+1}. Thus we showed that there is path

$$\mathbf{In} \xrightarrow{u_0} B_{i_0} \xrightarrow{u_1} B_{i_1} \xrightarrow{u_2} B_{i_2} \quad \cdots \quad B_{i_{s-1}} \xrightarrow{u_s} B_{i_s} \xrightarrow{u_{s+1}} \mathbf{Fin}$$

in $\mathcal{G}(M, \mathcal{F})$ from **In** to **Fin** of weight $u_0 t_{i_0} u_1 t_{i_1} u_2 t_{i_2} \cdots t_{i_{s-1}} u_s t_{i_s} u_{s+1} = t$. ◀

The next proposition provides a bound on the length of a path in $\mathcal{G}(M, \mathcal{F})$ with weight t.

▶ **Proposition 18.** *For any integer $t > 0$, if there is a path in $\mathcal{G}(M, \mathcal{F})$ from* **In** *to* **Fin** *of weight t, then there is such path of length at most $2m \log_2 t + 2m + \log_2 t$.*

Proof. Suppose P is a path in $\mathcal{G}(M, \mathcal{F})$ from **In** to **Fin** of weight t. Then the number of nodes and edges in P whose weight is greater than 1 or less than -1 is bounded by $\log_2 t$.

A *simple cycle* at node B_j is a closed path that starts and ends at B_j and in which no vertex appears twice except for B_j itself.

Note that if P contains a simple cycle of weight 1, then it can be removed from P without changing its weight. On the other hand, if P contains a simple cycle of weight -1, then removing such cycle will change the sign of the weight of P.

Let W_1 and W_2 be a successive pair of nodes or edges in P with weight different from ± 1. Then any node and edge that appears in P strictly between W_1 and W_2 has weight equal to ± 1. By the above observation we can remove all cycles of weight 1 that occur between W_1 and W_2 and leave at most one simple cycle of weight -1 between W_1 and W_2 in order to preserve the sign of the weight of P. So we can replace the original path from W_1 to W_2 by a new path with the same weight and length at most $2m$.

Recall there are at most $\log_2 t$ nodes and edges in P whose weight is different from ± 1. We now apply the above procedure to every pair W_1 and W_2 of successive nodes or edges in P whose weight is different from ± 1 including the cases when $W_1 = \mathbf{In}$ or $W_2 = \mathbf{Fin}$. There are at most $\log_2 t + 1$ such successive pairs. Therefore, we replace the whole path P with another path of the same weight and of length at most $2m(\log_2 t + 1) + \log_2 t$. Note that we added $\log_2 t$ in the end because every edge of weight different from ± 1 contributes 1 to the length of the path. ◀

Now we complete the proof of Theorem 13 using Propositions 17 and 18. Indeed, by Propositions 17 to decide whether $M \in \langle \mathcal{F} \rangle$, we need to check if there is a path in $\mathcal{G}(M, \mathcal{F})$ from **In** to **Fin** of weight t. By Propositions 18 the length of such path is bounded by $2m \log_2 t + 2m + \log_2 t$. Hence we can check all paths in $\mathcal{G}(M, \mathcal{F})$ of length up to $2m \log_2 t + 2m + \log_2 t$ to see if there is one with weight t. ◀

Conclusion and future work

The complexity of our algorithm is in EXPTIME. This is because a canonical word that represents a given matrix M has length exponential in the binary presentation of M. Hence the construction of regular languages in our proof takes exponential time. Moreover, the number of paths in $\mathcal{G}(M, \mathcal{F})$ of length up to $2m \log_2 t + 2m + \log_2 t$ is exponential in m.

In [4] it has been shown that the identity problem in $\mathrm{SL}(2,\mathbb{Z})$ is NP-complete. We would like to find out whether this construction can be combined with our result to show that the membership in $\mathrm{GL}(2,\mathbb{Z})$ extended by singular matrices is also NP-complete.

In our previous work [25] we proved that the membership problem is decidable for 2×2 nonsingular integer matrices. In this paper we considered matrices with determinants $0, \pm 1$. So, the next natural step will be to study the decidability of the membership problem for all 2×2 integer matrices, i.e. both singular and nonsingular ones.

References

1 László Babai, Robert Beals, Jin-yi Cai, Gábor Ivanyos, and Eugene M. Luks. Multiplicative equations over commuting matrices. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '96, pages 498–507, Philadelphia, PA, USA, 1996. Society for Industrial and Applied Mathematics.

2 Paul Bell and Igor Potapov. On undecidability bounds for matrix decision problems. *Theoretical Computer Science*, 391(1-2):3–13, 2008.

3 Paul C. Bell, Mika Hirvensalo, and Igor Potapov. Mortality for 2x2 matrices is NP-hard. In Branislav Rovan, Vladimiro Sassone, and Peter Widmayer, editors, *Mathematical Foundations of Computer Science 2012*, volume 7464 of *Lecture Notes in Computer Science*, pages 148–159. Springer Berlin Heidelberg, 2012.

4 Paul C. Bell, Mika Hirvensalo, and Igor Potapov. The identity problem for matrix semigroups in $\mathrm{SL}_2(\mathbb{Z})$ is NP-complete. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 187–206, 2017. URL: `http://dx.doi.org/10.1137/1.9781611974782.13`, `doi:10.1137/1.9781611974782.13`.

5 Paul C. Bell and Igor Potapov. On the undecidability of the identity correspondence problem and its applications for word and matrix semigroups. *Int. J. Found. Comput. Sci.*, 21(6):963–978, 2010.

6 Paul C. Bell and Igor Potapov. On the computational complexity of matrix semigroup problems. *Fundam. Inf.*, 116(1-4):1–13, January 2012.

7 Vincent D. Blondel, Emmanuel Jeandel, Pascal Koiran, and Natacha Portier. Decidable and undecidable problems about quantum automata. *SIAM J. Comput.*, 34(6):1464–1473, June 2005.

8 Julien Cassaigne, Vesa Halava, Tero Harju, and François Nicolas. Tighter undecidability bounds for matrix mortality, zero-in-the-corner problems, and more. *CoRR*, abs/1404.0644, 2014.

9 Christian Choffrut and Juhani Karhumaki. Some decision problems on integer matrices. *RAIRO-Theor. Inf. Appl.*, 39(1):125–131, 2005.

10 J. Esparza, A. Finkel, and R. Mayr. On the verification of broadcast protocols. In *Logic in Computer Science, 1999. Proceedings. 14th Symposium on*, pages 352–359, 1999.

11 Esther Galby, Joël Ouaknine, and James Worrell. On Matrix Powering in Low Dimensions. In Ernst W. Mayr and Nicolas Ollinger, editors, *32nd International Symposium on Theoretical Aspects of Computer Science (STACS 2015)*, volume 30 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 329–340, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

12 Yuri Gurevich and Paul Schupp. Membership problem for the modular group. *SIAM J. Comput.*, 37(2):425–459, May 2007.

13 Vesa Halava, Tero Harju, Mika Hirvensalo, and Juhani Karhumaki. Skolem's problem - on the border between decidability and undecidability. Technical Report 683, Turku Centre for Computer Science, 2005.

14 R. Kannan and R. J. Lipton. Polynomial-time algorithm for the orbit problem. *J. ACM*, 33(4):808–821, August 1986. URL: `http://doi.acm.org/10.1145/6490.6496`, `doi:10.1145/6490.6496`.

15 Ravindran Kannan and Achim Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM J. Comput.*, 8(4):499–507, 1979.

16 Alexei Lisitsa and Igor Potapov. Membership and reachability problems for row-monomial transformations. In *Mathematical Foundations of Computer Science 2004, 29th International Symposium, MFCS 2004, Prague, Czech Republic, August 22-27, 2004, Proceedings*, pages 623–634, 2004.

17 Roger C. Lyndon and Paul E. Schupp. *Combinatorial group theory*. Springer-Verlag, Berlin-New York, 1977. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 89.

18 Wilhelm Magnus, Abraham Karrass, and Donald Solitar. *Combinatorial group theory*. Dover Publications, Inc., New York, revised edition, 1976. Presentations of groups in terms of generators and relations.

19 C. Nuccio and Emanuele Rodaro. Mortality problem for 2×2 integer matrices. In *SOFSEM 2008: Theory and Practice of Computer Science, 34th Conference on Current Trends in Theory and Practice of Computer Science, Nový Smokovec, Slovakia, January 19-25, 2008, Proceedings*, pages 400–405, 2008. URL: `http://dx.doi.org/10.1007/978-3-540-77566-9_34`.

20 Joël Ouaknine, João Sousa Pinto, and James Worrell. On termination of integer linear loops. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, pages 957–969. SIAM, 2015.

21 Joël Ouaknine and James Worrell. On the positivity problem for simple linear recurrence sequences,. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 318–329, 2014.

22 Joël Ouaknine and James Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 330–341, 2014.

23 M. S. Paterson. Unsolvability in 3×3 matrices. *Studies in Applied Mathematics*, 49(1):pp.105–107, 1970.

24 Igor Potapov and Pavel Semukhin. Vector reachability problem in SL(2,Z). In *41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016, August 22-26, 2016 - Kraków, Poland*, pages 84:1–84:14, 2016. URL: `http://dx.doi.org/10.4230/LIPIcs.MFCS.2016.84`, `doi:10.4230/LIPIcs.MFCS.2016.84`.

25 Igor Potapov and Pavel Semukhin. Decidability of the membership problem for 2×2 integer matrices. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 170–186, 2017. URL: `http://dx.doi.org/10.1137/1.9781611974782.12`.

26 Robert A. Rankin. *Modular forms and functions*. Cambridge University Press, Cambridge-New York-Melbourne, 1977.